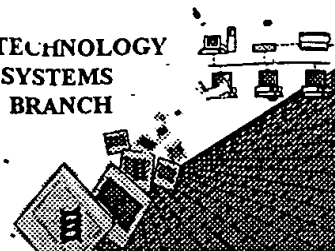BIOTECHNOLOGY
SYSTEMS
BRANCH

RAW SEQUENCE LISTING
ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/831, 426 B **RECEIVED**

Source: OIPE

Date Processed by STIC: 2/22/2002

MAR 1 4 2002

TECH CENTER 1600/2900

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY
FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703- 308-4216.
PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)
PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER VERSION 3.1 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:
http://www.uspto.gov/web/offices/pac/checker

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.
Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.
Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<http://www.uspto.gov/ebc/efs/downloads/documents.htm> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202
3. Hand Carry directly to:
   U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7th Floor, Examiner Name, Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202
        Or
   U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 01/29/2002

OIPE

Does Not Comply
Corrected Diskette Needed

```
 4 <110> APPLICANT: Hoechst Marion Roussel
 5        Bordon-Pallier, F.
 6        Rocher, C.
 8 <120> TITLE OF INVENTION: Human htFIIIA gene and coded htFIIIA protein
10 <130> FILE REFERENCE: 146.1364
12 <140> CURRENT APPLICATION NUMBER: US 09/831,426B
13 <141> CURRENT FILING DATE: 2001-05-08
15 <160> NUMBER OF SEQ ID NOS: 10
17 <170> SOFTWARE: PatentIn Vers. 2.0
```

RECEIVED

MAR 1 4 2002

TECH CENTER 1600/2900

ERRORED SEQUENCES

```
19 <210> SEQ ID NO: 1
20 <211> LENGTH: 1273
21 <212> TYPE: DNA
22 <213> ORGANISM: Human
24 <220> FEATURE:
25 <221> NAME/KEY: CDS
26 <222> LOCATION: (176)..(1270)
```

number for numeric identifier <400> should match numeric identifier <210>.

```
E--> 28 <400> SEQUENCE: (0)
```

The type of errors shown exist throughout the Sequence Listing. Please check subsequent sequences for similar errors.

```
29 atgcgcagca gcggcgccga cgcggggcgg tgcctggtga ccgcgcgcgc tcccggaagt
31 gtgccggcgt cgcgcgaagg ttcagcaggg agccgtgggc cgggcgcgcc ggttcccggc
33 acgtgtctcg gcacgtggca gcgcgcctgg ccctgggctt ggaggcgccg cgcc ctg
34                                                               Met
35                                                                1
37 gat ccg ccg gcc gtg gtc gcc gag tcg gtg tcg tcc ttg acc atc gcc   226
38 Asp Pro Pro Ala Val Val Ala Glu Ser Val Ser Ser Leu Thr Ile Ala
39              5                   10                  15
41 gac gcg ttc att gca gcc ggc gag agc tca gct ccg acc ccg ccg cgc   274
42 Asp Ala Phe Ile Ala Ala Gly Glu Ser Ser Ala Pro Thr Pro Pro Arg
43          20                  25                  30
45 ccc gcg ctt ccc agg agg ttc atc tgc tcc ttc cct gac tgc agc gcc   322
46 Pro Ala Leu Pro Arg Arg Phe Ile Cys Ser Phe Pro Asp Cys Ser Ala
47      35                  40                  45
49 aat tac agc aaa gcc tgg aag ctt gac gcg cac ctg tgc aag cac acg   370
50 Asn Tyr Ser Lys Ala Trp Lys Leu Asp Ala His Leu Cys Lys His Thr
51  50                  55                  60                  65
53 ggg gag aga cca ttt gtt tgt gac tat gaa ggg tgt ggc aag gcc ttc   418
54 Gly Glu Arg Pro Phe Val Cys Asp Tyr Glu Gly Cys Gly Lys Ala Phe
55                  70                  75                  80
57 atc agg gac tac cat ctg agc cgc cac att ctg act cac aca gga gaa   466
58 Ile Arg Asp Tyr His Leu Ser Arg His Ile Leu Thr His Thr Gly Glu
```

```
59                  85                    90                    95
61 aag ccg ttt gtt tgt gca gcc act ggc tgt gat caa aaa ttc aac aca    514
62 Lys Pro Phe Val Cys Ala Ala Thr Gly Cys Asp Gln Lys Phe Asn Thr
63          100                   105                   110
65 aaa tca aac ttg aag aaa cat ttt gaa cgc aaa cat gaa aat caa caa    562
66 Lys Ser Asn Leu Lys Lys His Phe Glu Arg Lys His Glu Asn Gln Gln
67      115                   120                   125
69 aaa caa tat ata tgc agt ttt gaa gac tgt aag aag acc ttt aag aaa    610
70 Lys Gln Tyr Ile Cys Ser Phe Glu Asp Cys Lys Lys Thr Phe Lys Lys
71 130                   135                   140                   145
73 cat cag cag ctg aaa atc cat cag tgc cag cat acc aat gaa cct cta    658
74 His Gln Gln Leu Lys Ile His Gln Cys Gln His Thr Asn Glu Pro Leu
75                   150                   155                   160
77 ttc aag tgt acc cag gaa gga tgt ggg aaa cac ttt gca tca ccc agc    706
78 Phe Lys Cys Thr Gln Glu Gly Cys Gly Lys His Phe Ala Ser Pro Ser
79                   165                   170                   175
81 aag ctg aaa cga cat gcc aag gcc cac gag ggc tat gta tgt caa aaa    754
82 Lys Leu Lys Arg His Ala Lys Ala His Glu Gly Tyr Val Cys Gln Lys
83          180                   185                   190
85 gga tgt tcc ttt gtg gca aaa aca tgg acg gaa ctt ctg aaa cat gtg    802
86 Gly Cys Ser Phe Val Ala Lys Thr Trp Thr Glu Leu Leu Lys His Val
87      195                   200                   205
89 aga gaa acc cat aaa gag gaa ata cta tgt gaa gta tgc cgg aaa aca    850
90 Arg Glu Thr His Lys Glu Glu Ile Leu Cys Glu Val Cys Arg Lys Thr
91 210                   215                   220                   225
93 ttt aaa cgc aaa gat tac ctt aag caa cac atg aaa act cat gcc cca    898
94 Phe Lys Arg Lys Asp Tyr Leu Lys Gln His Met Lys Thr His Ala Pro
95                   230                   235                   240
97 gaa agg gat gta tgt cgc tgt cca aga gaa ggc tgt gga aga acc tat    946
98 Glu Arg Asp Val Cys Arg Cys Pro Arg Glu Gly Cys Gly Arg Thr Tyr
99                   245                   250                   255
101 act act gtg ttt aat ctc caa agc cat atc ctc tcc ttc cat gag gaa    994
102 Thr Thr Val Phe Asn Leu Gln Ser His Ile Leu Ser Phe His Glu Glu
103          260                   265                   270
105 agc cgc cct ttt gtg tgt gaa cat gct ggc tgt ggc aaa aca ttt gca   1042
106 Ser Arg Pro Phe Val Cys Glu His Ala Gly Cys Gly Lys Thr Phe Ala
107      275                   280                   285
109 atg aaa caa agt ctc act agg cat gct gtt gta cat gat cct gac aag   1090
110 Met Lys Gln Ser Leu Thr Arg His Ala Val Val His Asp Pro Asp Lys
111 290                   295                   300                   305
113 aag aaa atg aag ctc aaa gtc aaa aaa tct cgt gaa aaa cgg agt ttg   1138
114 Lys Lys Met Lys Leu Lys Val Lys Lys Ser Arg Glu Lys Arg Ser Leu
115                   310                   315                   320
117 gcc tct cat ctc agt gga tat atc cct ccc aaa agg aaa caa ggg caa   1186
118 Ala Ser His Leu Ser Gly Tyr Ile Pro Pro Lys Arg Lys Gln Gly Gln
119                   325                   330                   335
121 ggc tta tct ttg tgt caa aac gga gag tca ccc aac tgt gtg gaa gac   1234
122 Gly Leu Ser Leu Cys Gln Asn Gly Glu Ser Pro Asn Cys Val Glu Asp
123          340                   345                   350
```

```
125 aag atg ctc tcg aca gtt gca gta ctt acc ctt ggc taa              1273
126 Lys Met Leu Ser Thr Val Ala Val Leu Thr Leu Gly
127     355             360             365
130 <210> SEQ ID NO: 2
131 <211> LENGTH: 365
132 <212> TYPE: PRT
133 <213> ORGANISM: Human          — see page 1
E--> 135 <400> SEQUENCE: (0)
136 Met Asp Pro Pro Ala Val Val Ala Glu Ser Val Ser Ser Leu Thr Ile
137   1             5               10              15
139 Ala Asp Ala Phe Ile Ala Ala Gly Glu Ser Ser Ala Pro Thr Pro Pro
140             20              25              30
142 Arg Pro Ala Leu Pro Arg Arg Phe Ile Cys Ser Phe Pro Asp Cys Ser
143         35              40              45
145 Ala Asn Tyr Ser Lys Ala Trp Lys Leu Asp Ala His Leu Cys Lys His
146     50              55              60
148 Thr Gly Glu Arg Pro Phe Val Cys Asp Tyr Glu Gly Cys Gly Lys Ala
149 65              70              75              80
151 Phe Ile Arg Asp Tyr His Leu Ser Arg His Ile Leu Thr His Thr Gly
152             85              90              95
154 Glu Lys Pro Phe Val Cys Ala Ala Thr Gly Cys Asp Gln Lys Phe Asn
155             100             105             110
157 Thr Lys Ser Asn Leu Lys Lys His Phe Glu Arg Lys His Glu Asn Gln
158         115             120             125
160 Gln Lys Gln Tyr Ile Cys Ser Phe Glu Asp Cys Lys Lys Thr Phe Lys
161     130             135             140
163 Lys His Gln Gln Leu Lys Ile His Gln Cys Gln His Thr Asn Glu Pro
164 145             150             155             160
166 Leu Phe Lys Cys Thr Gln Glu Gly Cys Gly Lys His Phe Ala Ser Pro
167             165             170             175
169 Ser Lys Leu Lys Arg His Ala Lys Ala His Glu Gly Tyr Val Cys Gln
170         180             185             190
172 Lys Gly Cys Ser Phe Val Ala Lys Thr Trp Thr Glu Leu Leu Lys His
173         195             200             205
175 Val Arg Glu Thr His Lys Glu Glu Ile Leu Cys Glu Val Cys Arg Lys
176     210             215             220
178 Thr Phe Lys Arg Lys Asp Tyr Leu Lys Gln His Met Lys Thr His Ala
179 225             230             235             240
181 Pro Glu Arg Asp Val Cys Arg Cys Pro Arg Glu Gly Cys Gly Arg Thr
182 .               245             250             .   255
184 Tyr Thr Thr Val Phe Asn Leu Gln Ser His Ile Leu Ser Phe His Glu
185             260             265             270
187 Glu Ser Arg Pro Phe Val Cys Glu His Ala Gly Cys Gly Lys Thr Phe
188         275             280             285
190 Ala Met Lys Gln Ser Leu Thr Arg His Ala Val Val His Asp Pro Asp
191         290             295             300
193 Lys Lys Lys Met Lys Leu Lys Val Lys Lys Ser Arg Glu Lys Arg Ser
194 305             310             315             320
196 Leu Ala Ser His Leu Ser Gly Tyr Ile Pro Pro Lys Arg Lys Gln Gly
```

```
197                 325                 330                 335
199 Gln Gly Leu Ser Leu Cys Gln Asn Gly Glu Ser Pro Asn Cys Val Glu
200                 340                 345                 350
202 Asp Lys Met Leu Ser Thr Val Ala Val Leu Thr Leu Gly
203         355                 360                 365
206 <210> SEQ ID NO: 3
207 <211> LENGTH: 1273
208 <212> TYPE: DNA
209 <213> ORGANISM: Human                    See page 1
211 <400> SEQUENCE: (0)
212 atgcgcagca gcggcgccga cgcggggcgg tgcctggtga ccgcgcgcgc tcccggaagt 60
214 gtgccggcgt cgcgcgaagg ttcagcaggg agccgtgggc cgggcgcgcc ggttcccggc 120
216 acgtgtctcg gcacgtggca gcgcgcctgg ccctgggctt ggaggcgccg gcgccctgga 180
218 tccgccggcc gtggtcgccg agtcggtgtc gtccttgacc atcgccgacg cgttcattgc 240
220 agccggcgag agctcagctc cgaccccgcc gcgccccgcg cttcccagga ggttcatctg 300
222 ctccttccct gactgcagcg ccaattacag caaagcctgg aagcttgacg cgcacctgtg 360
224 caagcacacg ggggagagac catttgtttg tgactatgaa gggtgtggca aggccttcat 420
226 cagggactac catctgagcc gccacattct gactcacaca ggagaaaagc cgtttgtttg 480
228 tgcagccact ggctgtgatc aaaaattcaa cacaaaatca aacttgaaga aacattttga 540
230 acgcaaacat gaaaatcaac aaaaacaata tatatgcagt tttgaagact gtaagaagac 600
232 ctttaagaaa catcagcagc tgaaaatcca tcagtgccag cataccaatg aacctctatt 660
234 caagtgtacc caggaaggat gtgggaaaca ctttgcatca cccagcaagc tgaaacgaca 720
236 tgccaaggcc cacgagggct atgtatgtca aaaaggatgt tcctttgtgg caaaaacatg 780
238 gacggaactt ctgaaacatg tgagagaaac ccataaagag gaaatactat gtgaagtatg 840
240 ccggaaaaca tttaaacgca agattacct taagcaacac atgaaaactc atgccccaga 900
242 aagggatgta tgtcgctgtc caagagaagg ctgtggaaga acctatacta ctgtgtttaa 960
244 tctccaaagc catatcctct ccttccatga ggaaagccgc ccttttgtgt gtgaacatgc 1020
246 tggctgtggc aaaacatttg caatgaaaca aagtctcact aggcatgctg ttgtacatga 1080
248 tcctgacaag aagaaaatga agctcaaagt caaaaaatct cgtgaaaaac ggagtttggc 1140
250 ctctcatctc agtggatata tccctcccaa aaggaaacaa gggcaaggct tatctttgtg 1200
252 tcaaaacgga gagtcaccca actgtgtgga agacaagatg ctctcgacag ttgcagtact 1260
254 tacccttggc taa                                                 1273
257 <210> SEQ ID NO: 4
258 <211> LENGTH: 1213
259 <212> TYPE: DNA
260 <213> ORGANISM: Human
262 <400> SEQUENCE: (0)
263 gtgccggcgc cgcgcgaagg ttcagcaggg agccgtgggc cgggcgcgcc ggttcccggc 60
265 acgtgtctcg gcacgtggca gcgcgcctgg ccctgggctt ggaggcgccg gcgccctgga 120
267 tccgccggcc gtggtcgccg agtcggtgtc gtccttgacc atcgccgacg cgttcattgc 180
269 agccggcgag agctcagctc cgaccccgcc gcgccccgcg cttcccagga ggttcatctg 240
271 ctccttccct gactgcagcg ccaattacag caaagcctgg aagcttgacg cgcacctgtg 300
273 caagcacacg ggggagagac catttgtttg tgactatgaa gggtgtggca aggccttcat 360
275 cagggactac catctgagcc gccacattct gactcacaca ggagaaaagc cgtttgtttg 420
277 tgcagccact ggctgtgatc aaaaattcaa cacaaaatca aacttgaaga aacattttga 480
279 acgcaaacat gaaaatcaac aaaaacaata tatatgcagt tttgaagact gtaagaagac 540
281 ctttaagaaa catcagcagc tgaaaatcca tcagtgccag cataccaatg aacctctatt 600
283 caagtgtacc caggaaggat gtgggaaaca ctttgcatca cccagcaagc tgaaacgaca 660
285 tgccaaggcc cacgagggct atgtatgtca aaaaggatgt tcctttgtgg caaaaacatg 720
```

Input Set : A:\9823seq.txt
Output Set: N:\CRF3\02222002\I831426B.raw

```
     287  gacggaactt ctgaaacatg tgagagaaac ccataaagag gaaatactat gtgaagtatg 780
     289  ccggaaaaca tttaaacgca aagattacct taagcaacac atgaaaactc atgccccaga 840
     291  aagggatgta tgtcgctgtc caagagaagg ctgtggaaga acctatacta ctgtgtttaa 900
     293  tctccaaagc catatcctct ccttccatga ggaaagccgc ccttttgtgt gtgaacatgc 960
     295  tggctgtggc aaaacatttg caatgaaaca aagtctcact aggcatgctg ttgtacatga 1020
     297  tcctgacaag aagaaaatga agctcaaagt caaaaaatct cgtgaaaaac ggagtttggc 1080
     299  ctctcatctc agtggatata tccctcccaa aaggaaacaa gggcaaggct tatctttgtg 1140
     301  tcaaaacgga gagtcaccca actgtgtgga agacaagatg ctctcgacag ttgcagtact 1200
     303  tacccttggc taa                                                  1213
     306  <210> SEQ ID NO: 5
     307  <211> LENGTH: 34
     308  <212> TYPE: DNA
     309  <213> ORGANISM: Human
E--> 311  <400> SEQUENCE: ⓞ
     312  cggggtacca aaaatgcgca gcagcggcgc cgac                           34
     315  <210> SEQ ID NO: 6
     316  <211> LENGTH: 21
     317  <212> TYPE: DNA
     318  <213> ORGANISM: Human
E--> 320  <400> SEQUENCE: ⓞ
     321  tccttccctg actgcagcgc c                                        21
     324  <210> SEQ ID NO: 7
     325  <211> LENGTH: 20
     326  <212> TYPE: DNA
     327  <213> ORGANISM: Human
E--> 329  <400> SEQUENCE: ⓞ ———— see page 1
     330  tgcacaggtg cgcgtcaagc                                          20
     333  <210> SEQ ID NO: 8
     334  <211> LENGTH: 20
     335  <212> TYPE: DNA
     336  <213> ORGANISM: Human
E--> 338  <400> SEQUENCE: ⓞ
     339  cacaaacaaa tggtctctcc                                          20
     342  <210> SEQ ID NO: 9
     343  <211> LENGTH: 30
     344  <212> TYPE: DNA
     345  <213> ORGANISM: Human
E--> 347  <400> SEQUENCE: ⓞ
     348  cggtctagat tagccaaggg taagtactgc                               30
     351  <210> SEQ ID NO: 10
     352  <211> LENGTH: 30
     353  <212> TYPE: DNA
     354  <213> ORGANISM: Human
E--> 356  <400> SEQUENCE: ⓞ
     357  cctcccgggg ccaagggtaa gtactgcaac                               30
E--> 364  ⑦
```

remove extra material at end of file.

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/831,426B

DATE: 02/22/2002
TIME: 15:16:40

Input Set : **A:\9823seq.txt**
Output Set: **N:\CRF3\02222002\I831426B.raw**

```
L:28 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQUENCE ID NOS:1 differs:0
L:135 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQUENCE ID NOS:2 differs:0
L:211 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQUENCE ID NOS:3 differs:0
L:262 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQUENCE ID NOS:4 differs:0
L:311 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQUENCE ID NOS:5 differs:0
L:320 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQUENCE ID NOS:6 differs:0
L:329 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQUENCE ID NOS:7 differs:0
L:338 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQUENCE ID NOS:8 differs:0
L:347 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQUENCE ID NOS:9 differs:0
L:356 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQUENCE ID NOS:10 differs:0
L:364 M:254 E: No. of Bases conflict, LENGTH:Input:7 Counted:30 SEQ:10
```